

Claims

1. A method for matching a spoken text, the method comprising:
receiving the spoken text from a voice recognition system
that converts a voice signal to the spoken text;
5 looking up for one or more words that match the spoken text,
wherein the one or more words are derived from an
identifier of a piece of information provided from a server
is coupled over a data network; and
10 retrieving the information either from the server or a local
database with respect to the identifier as a query to the
server or the database.
2. The method of claim 1, wherein the server is remotely located
and provides the information upon receiving a request including
15 the identifier.
3. The method of claim 2 further comprising:
generating the request to include the identifier; and
20 sending out the request over the data network.
4. The method of claim 3, wherein the data network includes one
of (i) the Internet, (ii) the Intranet, (iii) a wireless network, and
25 (iv) a network of a private and a public network.
5. The method of claim 4, wherein the voice signal is received from
a voice network and input to the voice recognition system.

6. The method of claim 5, wherein the voice network includes one or more of a public switched telephone network (PSTN) and a wireless network.
- 5
7. The method of claim 3, wherein the looking up for one or more words comprises:
 - receiving the identifier from the server, wherein the identifier contains more than one words;
 - 10 extracting one or more key words from the identifier; and
 - archiving the one or more key words in a local searching data set, wherein the local searching data set is remotely separated from the server.
- 15
8. The method of claim 7, wherein the extracting of the one or more key words from the identifier comprises discarding generic words from the identifier, wherein the generic words can be included in other identifiers.
- 20
9. The method of claim 7, wherein the extracting of the one or more key words from the identifier further comprises:
 - computing a histogram of the identifier; and
 - identifying the generic words and the key words.
- 25
10. The method of claim 1, wherein the retrieving of the information comprises obtaining the information from the local database if

the local database archives a copy of the information, otherwise from the server.

11. A method for matching a spoken text, the method comprising:

- 5 receiving a plurality of identifiers, each of the identifiers
 indicating a piece of information;
 identifying generic words and key words from the identifiers;
 and
 organizing the keywords in a structure that can facilitate
10 matching of the spoken text to one of the keywords.

12. The method of claim 11 further comprising storing the structure in a local database.

- 15 13. The method of claim 12, wherein the structure is based on a
 data structure designated from searching through with respect
 to a query.

14. The method of claim 13, wherein the structure is a tree
20 structure, each node of the tree structure being one of the
 keywords and associated with one or more of the key words that
 are collectively from one of the identifiers.

15. The method of claim 11, wherein each of the identifiers contains
25 one or more words and is selected from a group consisting of a
 title, a filename, a symbol, an IP address and a short article.

16. The method of claim 15, wherein the identifiers are supplied from a server providing information, each of pieces of the information being identified by one of the identifiers.
- 5 17. The method of claim 16, wherein the server is remotely located over a data network, and the identifiers are provided over the data network.
- 10 18. The method of claim 11, wherein the identifying of the generic words and the key words comprises:
 - computing a statistic measurement of the identifiers; wherein the statistic measurement indicating how often each of the generic words and the key words has respectively occurred in the identifiers; and
 - 15 classifying the generic words and the key words from the statistic measurement.
- 20 19. The method of claim 18, wherein the computing of the statistic measurement of the identifiers is computing a histogram of the identifiers.
- 25 20. The method of claim 19, wherein the classifying of the generic words and the key words comprises:
 - identifying any marginal words from the identifiers with respect to the histogram;

performing an linguistic analysis of the marginal words to
group the marginal words with either the generic words or
the key words.

5 21. A product including computer instructions to be executed in a
computing device, the product comprising:

program code for receiving the spoken text from a voice
recognition system that converts a voice signal to the
spoken text;

10 program code for looking up for one or more words that
match the spoken text, wherein the one or more words
are derived from an identifier of a piece of information
provided from a server is coupled over a data network;
and

15 program code for retrieving the information either from the
server or a local database with respect to the identifier as
a query to the server or the database.

22. The product of claim 21, wherein the server is remotely located
20 and provides the information upon receiving a request including
the identifier.

23. The product of claim 22 further comprising:
program code for generating the request to include the
25 identifier; and
program code for sending out the request over the data
network.

24. The product of claim 23, wherein the data network includes one of (i) the Internet, (ii) the Intranet, (iii) a wireless network, and (iv) a network of a private and a public network.
- 5
25. The product of claim 24, wherein the voice signal is received from a voice network and input to the voice recognition system.
26. The product of claim 25, wherein the voice network includes one or more of a public switched telephone network (PSTN) and a wireless network.
- 10
27. The product of claim 23, wherein the program code for looking up for one or more words comprises:
 - 15 program code for receiving the identifier from the server, wherein the identifier contains more than one words;
 - program code for extracting one or more key words from the identifier; and
 - program code for archiving the one or more key words in a local searching data set, wherein the local searching data set is remotely separated from the server.
- 20
28. The product of claim 27, wherein the program code for extracting the one or more key words from the identifier comprises discarding generic words from the identifier, wherein the generic words can be included in other identifiers.
- 25

29. The product of claim 27, wherein the program code for extracting the one or more key words from the identifier further comprises:

5

program code for computing a histogram of the identifier;
and
program code for identifying the generic words and the key words.

30. The product of claim 21, wherein the program code for retrieving the information comprises program code for obtaining the information from the local database if the local database archives a copy of the information, otherwise from the server.

10

31. A product including computer instructions to be executed in a computing device, the product comprising:

15

program code for receiving a plurality of identifiers, each of the identifiers indicating a piece of information;
program code for identifying generic words and key words from the identifiers; and

20

program code for organizing the keywords in a structure that can facilitate matching of the spoken text to one of the keywords.

32. The product of claim 31 further comprising program code for storing the structure in a local database.

25

33. The product of claim 32, wherein the structure is based on a data structure designated from searching through with respect to a query.
- 5 34. The product of claim 33, wherein the structure is a tree structure, each node of the tree structure being one of the keywords and associated with one or more of the key words that are collectively from one of the identifiers.
- 10 35. The product of claim 31, wherein each of the identifiers contains one or more words and is selected from a group consisting of a title, a filename, a symbol, an IP address and a short article.
- 15 36. The product of claim 35, wherein the identifiers are supplied from a server providing information, each of pieces of the information being identified by one of the identifiers.
- 20 37. The product of claim 36, wherein the server is remotely located over a data network, and the identifiers are provided over the data network.
- 25 38. The product of claim 31, wherein the program code for identifying of the generic words and the key words comprises:
 program code for computing a statistic measurement of the identifiers; wherein the statistic measurement indicating how often each of the generic words and the key words has respectively occurred in the identifiers; and

program code for classifying the generic words and the key words from the statistic measurement.

39. The product of claim 38, wherein the program code for computing the statistic measurement of the identifiers is computing a histogram of the identifiers.

40. The product of claim 39, wherein the program code for classifying the generic words and the key words comprises:
program code for identifying any marginal words from the identifiers with respect to the histogram;
program code for performing an linguistic analysis of the marginal words to group the marginal words with either the generic words or the key words.